Case Study: The Non-Small Cell Lung Cancer Symptom Assessment Questionnaire (NSCLC-SAQ)

Eighth Annual Patient-Reported Outcome Consortium Workshop

April 26 – 27, 2017 Bethesda, MD
Disclaimer

• The views and opinions expressed in the following slides are those of the individual presenters and should not be attributed to their respective organizations/companies, the U.S. Food and Drug Administration or the Critical Path Institute.

• These slides are the intellectual property of the individual presenters and are protected under the copyright laws of the United States of America and other countries. Used by permission. All rights reserved. All trademarks are the property of their respective owners.
Session Objectives

• To provide an overview of the development process of the *NSCLC-SAQ*
• To highlight the usefulness and successes of the mixed methods approach in its development
Session Participants

Moderator
– Kendra DeBusk, PhD – Principal Outcomes Research Scientist, Genentech, Inc.

Presenters
– Astra Liepa, PharmD – Principal Research Scientist, Eli Lilly and Company
– Kelly McCarrier, PhD, MPH – Senior Research Scientist, Health Research Associates, Inc.
– Don Bushnell, MA – Associate Director, Health Research Associates, Inc.
– Thomas Atkinson, PhD – Head, Behavioral Research Methods Core Facility, MSKCC
– Kendra DeBusk, PhD – Principal Outcomes Research Scientist, Genentech, Inc.

Panelists
– Paul Kluetz, MD – Acting Deputy Director, Office of Hematology and Oncology Products, FDA
– Stephen Joel Coons, PhD – Executive Director, PRO Consortium, C-Path
Session Outline

• Introduction
• Qualitative Development Steps
• Quantitative Pilot Study
• Challenges at the Site Level
• Summary of Regulatory Interactions
• Ongoing and Future Activities
• Panel Discussion and Q & A
Introduction

Astra Liepa, Principal Research Scientist, Eli Lilly and Company
NSCLC-SAQ Development Team

• NSCLC Working Group

• Co-Chairs
  • Astra Liepa
  • Kendra DeBusk

• Sponsors
  – AbbVie
  – AstraZeneca
  – Boehringer Ingelheim
  – Bristol-Myers Squibb
  – Eli Lilly and Company
  – EMD Serono
  – Genentech, Inc.
  – Janssen Global Services
  – Merck Sharp & Dohme
  – Novartis Pharmaceuticals
NSCLC-SAQ Development Team

• Expert Panel Members
  • Ethan Basch, MD, University of North Carolina at Chapel Hill
  • David Cella, PhD, Northwestern University
  • Shirish Gadgeel, MD, Karmanos Cancer Center
  • Richard J. Gralla, MD, Albert Einstein College of Medicine
  • Donald L. Patrick, PhD, University of Washington
  • Suresh Ramalingam, MD, Emory University

• Health Research Associates
  • Mona Martin, RN, MPA, Executive Director
  • Donald Bushnell, MA, Associate Director
  • Kelly McCarrier, PhD, Senior Research Scientist
  • Thomas Atkinson, PhD, Memorial Sloan Kettering Cancer Center

• Critical Path Institute’s PRO Consortium
  • Stephen Joel Coons, PhD, Executive Director
  • Sonya Eremenco, MA, Associate Director
  • Sarah Mann, MBA, PMP, Senior Project Manager
  • Theresa “T” Griffey, MBA, PMP, Senior Project Manager
Working Group Goal

• Goal of the PRO Consortium’s NSCLC Working Group
  • To achieve FDA qualification of a PRO instrument to assess symptoms of advanced NSCLC that is fit for purpose to support efficacy endpoints in clinical trials
    • Specific Context of Use: Advanced NSCLC (Stage IIIB/IV, Performance Status 0 to 2, any line of therapy)

• Rationale for the NSCLC Working Group (WG)
  • PRO Consortium members and FDA advisors identified NSCLC as a priority area.
  • It was unclear at the outset whether any existing PRO instruments were ‘fit for purpose’ as an efficacy endpoint in NSCLC clinical trials
    • While existing PRO measures are available for the assessment of NSCLC symptoms, none had been shown to meet FDA’s current expectations for supporting label claims.
Why a New Measure?

• There was an apparent lack of a PRO instrument developed in accordance with the FDA PRO Guidance for use in clinical trials
• The Working Group agreed that de novo instrument development (with literature review, concept elicitation and cognitive interviews, and quantitative evaluation) would provide the highest likelihood of success in the COA qualification process.
  • All evidence of patient involvement would be readily available to FDA reviewers
  • Specific concerns raised by the FDA could be addressed through the development process
• De novo development of a brief measure of core disease symptoms
  • Did not include treatment-related symptoms or impacts
Qualitative Development Steps

Kelly McCarrier, Senior Research Scientist, HRA, Inc.
Flow of Qualitative Development Steps

1. Conduct Literature and PRO Instrument review
2. Develop study protocol, interview guide and study forms
3. Concept Elicitation (CE) Interviews
4. Review CE data, Complete item generation process, Format draft NSCLC-SAQ
5. Program ePRO (Tablet) version of draft NSCLC-SAQ
6. Evaluate & refine NSCLC-SAQ through cognitive interviews, Translatability Assessment, Electronic Implementation Assessment
7. Qualitative evidence dossier submitted to FDA
8. Instrument Refinement and Launch of Quantitative Pilot

Timeline:
- 12/2012 to 04/2013
- 04/2013 to 06/2013
- 06/2013 to 07/2013
- 12/2013 to 03/2014
- 03/2014
- 04/2015
Concept-Focused Literature Review

• Objective/Approach:
  • The purpose of this review was to identify NSCLC symptom and impact related concepts identified in previously-published qualitative research.
  • Searches conducted in Pubmed, EMBASE, PsycINFO and Google Scholar (limited to English articles published between 1992 and 2012)

• Findings:
  • Between primary and secondary searches, 893 abstracts reviewed; 27 articles retained for initial full review; and final review included 19 articles
  • 32 sign/symptom concepts were identified (4 symptom areas: Physical, Gastrointestinal distress, Respiratory distress and Malaise)
  • 30 impacts concepts were identified (4 impact areas: QOL, Physical functioning, Relations with others and Psychological)
  • The concepts found in the review influenced the development of the CE interview guide.
In Instrument Review:

**Objective/Approach:**
- To summarize key characteristics of a selection of existing instruments to support decision-making in defining the conceptual (disease) model, identifying preliminary key concepts, and helping develop the qualitative interview guide.
- Searches conducted in Pubmed, PsycINFO, EMBASE, ProQolid, and ISPOR (limited to English articles published between 2002 and 2012).

**Findings:**
- Between primary and secondary searches, 821 abstracts reviewed; 191 articles retained for initial full review.
- Identified 45 PRO instruments.
- 11 selected for in-depth review (CARES, CDS, EORTC QLQ-LC13, FACIT-F, FACT-L, FLSI, LCSS, MDASI, MRC, RSCL, SDS).
- The search included the use of qualitative evidence of concept relevance during the development of the measure, as well as pattern of use, and satisfactory psychometric performance characteristics.
Concept Elicitation (CE) Interviews

• Objective:
  • To identify and document symptom concepts that are relevant and important to patients with NSCLC in order to support PRO measure development for NSCLC clinical trials

• Methods:
  • N=51 individual, face-to-face qualitative CE interviews
  • Patients recruited from 6 clinics in U.S.
  • 60-90 minute interviews; following a semi-structured CE interview guide (including open-ended items, day-reconstruction, probing, and rating exercises)
Key Eligibility Criteria

• Inclusion Criteria:
  • Adults with diagnosis of Stage I-IV NSCLC
  • ECOG Performance Status of 0-2
  • Diagnosed with Stage I or II and naïve to treatment for NSCLC; or
  • Diagnosed with Stage III or IV cancer and naïve to treatment for NSCLC or had recovered from any prior treatment-related toxicity/adverse events to CTCAE v4.03 grade 1 (mild) or better

• Exclusions:
  • Current / past history of personality disorder, bipolar disorder, schizophrenia or other psychotic disorder, OCD, PTSD, mental retardation, organic mental disorders, or mental disorders due to a general medical condition
  • Recent (12-month) history of clinically significant drug or alcohol abuse or dependence, excluding nicotine
  • Currently enrolled in any investigational device, drug, or biologics product study
### Demographic Characteristics

<table>
<thead>
<tr>
<th>Demographic Characteristics</th>
<th>Total subject N=51 (100%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
</tr>
<tr>
<td>- Mean (SD)</td>
<td>64.9 (11.2)</td>
</tr>
<tr>
<td>- Median</td>
<td>66</td>
</tr>
<tr>
<td>- Range</td>
<td>46-86</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
</tr>
<tr>
<td>- Male</td>
<td>25 (49%)</td>
</tr>
<tr>
<td>- Female</td>
<td>26 (51%)</td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
</tr>
<tr>
<td>- Hispanic, Latino, or Spanish</td>
<td>5 (10%)</td>
</tr>
<tr>
<td>- Not Hispanic or Latino</td>
<td>46 (90%)</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
</tr>
<tr>
<td>- White</td>
<td>38 (75%)</td>
</tr>
<tr>
<td>- Black/African American</td>
<td>8 (16%)</td>
</tr>
<tr>
<td>- Asian</td>
<td>2 (4%)</td>
</tr>
<tr>
<td>- Other</td>
<td>3 (6%)</td>
</tr>
<tr>
<td><strong>Highest level of education completed</strong></td>
<td></td>
</tr>
<tr>
<td>- Less than High School</td>
<td>3 (6%)</td>
</tr>
<tr>
<td>- High School</td>
<td>25 (49%)</td>
</tr>
<tr>
<td>- Some College</td>
<td>13 (26%)</td>
</tr>
<tr>
<td>- Bachelor’s Degree</td>
<td>3 (6%)</td>
</tr>
<tr>
<td>- Graduate or Professional School</td>
<td>7 (14%)</td>
</tr>
</tbody>
</table>
## Clinical Characteristics

<table>
<thead>
<tr>
<th>Clinical Characteristic</th>
<th>Total Subjects N=51 (100%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diagnosis of Comorbid COPD</strong></td>
<td></td>
</tr>
<tr>
<td>- Yes</td>
<td>18 (35%)</td>
</tr>
<tr>
<td>- No</td>
<td>33 (65%)</td>
</tr>
<tr>
<td><strong>Stage at initial NSCLC diagnosis</strong></td>
<td></td>
</tr>
<tr>
<td>- Stage I</td>
<td>6 (12%)</td>
</tr>
<tr>
<td>- Stage II</td>
<td>1 (2%)</td>
</tr>
<tr>
<td>- Stage III</td>
<td>25 (49%)</td>
</tr>
<tr>
<td>- Stage IV</td>
<td>19 (37%)</td>
</tr>
<tr>
<td><strong>Current line of NSCLC treatment</strong></td>
<td></td>
</tr>
<tr>
<td>- Early</td>
<td>19 (37%)</td>
</tr>
<tr>
<td>- 1&lt;sup&gt;st&lt;/sup&gt; line advanced/metastatic</td>
<td>18 (35%)</td>
</tr>
<tr>
<td>- 2&lt;sup&gt;nd&lt;/sup&gt; line advanced/metastatic</td>
<td>9 (18%)</td>
</tr>
<tr>
<td>- 3&lt;sup&gt;rd&lt;/sup&gt; line advanced/metastatic</td>
<td>3 (6%)</td>
</tr>
<tr>
<td>- Other: Observation, Subsequent</td>
<td>2 (4%)</td>
</tr>
<tr>
<td><strong>Current ECOG performance status</strong></td>
<td></td>
</tr>
<tr>
<td>- ECOG 0</td>
<td>17 (33%)</td>
</tr>
<tr>
<td>- ECOG 1</td>
<td>24 (47%)</td>
</tr>
<tr>
<td>- ECOG 2</td>
<td>10 (20%)</td>
</tr>
</tbody>
</table>
CE Findings: Saturation

• Within 51 CE transcripts, 1897 patient-based symptom expressions were coded

• Expressions were grouped into 43 distinct symptom concepts

• Evidence of concept saturation was observed:
  • 93% of concepts were identified within the first 9 interviews (18% of transcripts)
  • No new concepts appeared within the final 47% of transcripts

<table>
<thead>
<tr>
<th>CE Transcript Group</th>
<th>Group 1 (n=9)</th>
<th>Group 2 (n=9)</th>
<th>Group 3 (n=9)</th>
<th>Group 4 (n=8)</th>
<th>Group 5 (n=8)</th>
<th>Group 6 (n=8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of newly-coded concepts identified</td>
<td>40</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Percent of total (n=43) concepts coded</td>
<td>93%</td>
<td>5%</td>
<td>2%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>
• Key features examined for expressed concepts:
  • Frequency of mention (by subject, across all coded expressions)
  • Spontaneous mentions (vs. probed)
  • Subject ratings of bother and severity
    • 0 to 10 numeric rating scale (NRS) exercise for each expressed symptom
  • Subject ratings of meaningful attributes
    • Frequency, Severity, Duration
CE Findings: Concept Predominance

Among the 43 symptom concepts described by patients, the most frequently-expressed were:

- “Coughing” 206 expressions by 51 (100%) of interviewed subjects
- “Tiredness” 172 expressions by 39 (77%) subjects
- “Shortness of Breath” 149 expressions by 35 (67%) subjects
- “Poor Appetite” 92 expressions by 26 (51%) subjects
- “Difficulty Breathing” 91 expressions by 21 (41%) subjects
- “Chest Pain” 67 expressions by 17 (33%) subjects
- “General Pain” 59 expressions by 18 (35%) subjects
- “Low Energy” 41 expressions by 15 (29%) subjects
CE Findings: Spontaneous Mentions

• Among the 43 symptom concepts, the following were most frequently expressed spontaneously (as opposed to in response to a probe by the interviewer)
  • “Coughing” (expressed spontaneously by 29 subjects)
  • “Shortness of Breath” (22)
  • “Chest Pain” (14)
  • “Tiredness” (12)
  • “Less Appetite” (11)
  • “Difficulty Breathing” (9)
  • “General Pain” (8)
Item Generation Process

• Attended by WG, Expert Panel, C-Path, and HRA

• Reviewed key evidence from:
  • Literature and Instrument Review
  • CE Interview Findings
  • Expert Input

• Consensus reduced 43 symptom concepts to 9 that were targeted for evaluation in cognitive interviews
Conceptual framework following item generation for draft *NSCLC-SAQ*

**NSCLC Symptoms**

- Cough
  - Coughing up blood
- Hemoptysis
- Pain
  - Pain in chest
- Dyspnea
  - Shortness of breath
  - Difficulty breathing
- Fatigue
  - Lack of energy
- Appetite
  - Tires easily
  - Appetite
## Initial Draft **NSCLC-SAQ**

- Draft **NSCLC-SAQ** formatted for cognitive interviews
  - Stem wording drafted for each concept/item using patient language from CE data
  - 7-day recall period
  - Tested both a 5-level verbal rating scale (VRS) and a 0 to 10 NRS of symptom intensity (3 items) and frequency (6 items)

<table>
<thead>
<tr>
<th>Cough</th>
<th>Coughing</th>
<th>Over the past 7 days, how much did you cough?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Not at All</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A Little Bit</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quite a Bit</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Very Much</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cough</th>
<th>Coughing Up Blood (Hemoptysis)</th>
<th>Over the past 7 days, how often did you cough up blood?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Never</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rarely</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sometimes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Often</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Always</td>
</tr>
</tbody>
</table>

Please rate your cough over the last 7 days.  
0 – No Cough  
10 – As bad as you can imagine

Please rate the amount of blood you coughed up over the last 7 days.  
0 – No Coughing up Blood  
10 – Coughing blood as bad as you can imagine
Cognitive Interviews

• Objectives:
  • To evaluate clarity, comprehension, and relevance of the draft NSCLC-SAQ items, instructions, and response options
  • To assess comparability of paper and ePRO formats

• Methods:
  • N=20 individual, face-to-face cognitive interviews (~ 60 min. each)
  • Identical recruitment process as concept elicitation phase
  • Conducted in three waves:
    • Wave 1: n=4 interviews using paper format of draft NSCLC-SAQ
    • Wave 2: n=10 using ePRO and paper formats
    • Wave 3: n=6 using ePRO and paper formats
Translatability and Electronic Implementation Assessments

• Translatability Assessment
  • Objective: To identify potential difficulty in maintaining conceptual equivalence in translations
  • Methods:
    • Examined *NSCLC-SAQ* in Chinese, Hindi, Japanese, Russian, and Spanish
    • Linguistics consultants rated each element (items and instructions) from 1 (no difficulty) to 5 (extremely difficult)

• Electronic Implementation Assessment
  • Objective: To identify potential difficulty in implementing across full range of ePRO formats
  • Methods:
    • Conducted by members of the ePRO Consortium
    • Provided suggestions to maximize equivalence between paper and ePRO (generally) as well as across different ePRO formats
Key NSCLC-SAQ Modifications

• Wave 1
  • 5-level VRS retained over the 0 to 10 NRS format
  • Hemoptysis item removed

• Wave 2 and 3
  • Evidence of comparability between paper and ePRO format was observed
  • Severity/intensity-focused items (cough, chest pain, non-chest pain) reworded to assess peak (“worst”) intensity
  • The two dyspnea-focused items were combined into “short of breath during usual activities”
  • The appetite-focused item was reworded from assessing “good appetite” to “poor appetite” to keep directional consistency with response scales of other items in the measure
Conceptual framework for draft NSCLC-SAQ

NSCLC Symptoms:
- Cough
- Pain
- Dyspnea
- Fatigue
- Appetite

Cough
- Pain in chest
- Pain in areas other than your chest

Dyspnea
- Shortness of breath

Fatigue
- Lack of energy

Appetite
- Tires easily
- Poor appetite

After Cognitive Interviews and FDA Feedback (From Annual PRO Consortium Workshop, April 2015)
For Further Information...

The initial qualitative development research has recently been published in Clinical Therapeutics:

Quantitative Pilot Study

Don Bushnell, Associate Director, HRA, Inc.
The overall aim was to evaluate, and refine if warranted, the draft version of the NSCLC-SAQ. The main objectives were to:

- assess item performance using standard classical test theory item reduction methods (including evaluation of missing data, ceiling/floor effects, item-to-item correlations, item-to-total correlations, factor analysis and internal consistency reliability),
- assess item performance using Rasch measurement theory (RMT),
- examine the measurement model and derive a provisional NSCLC-SAQ scoring function,
- estimate one-week test-retest reliability, and
- examine construct validity.
Quantitative Study Design

• A sample of 150 subjects with clinically-diagnosed NSCLC was targeted from clinical sites to participate by completing a questionnaire battery twice (7 to 10 days apart) via a touchscreen tablet computer at the clinical sites.

• Each site recruited eligible subjects with Stage IIIB or IV NSCLC. Quotas were set for:
  • comorbid COPD,
  • line of treatment,
  • ECOG performance status, and
  • NSCLC stage.

• No change in treatment was required in order to participate in this study.
# Quantitative Data Collection Schema

## Timing of Study Assessments

<table>
<thead>
<tr>
<th>Day</th>
<th>Study Visit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>7-10†</td>
<td>2</td>
</tr>
</tbody>
</table>

---

### Completed by Clinic Staff Prior to Subject Assessments

- Screening Form *<completed prior to enrollment>*
- Clinician Information Form (CTCAE, ECOG performance status, clinical characteristics) *<completed within 24-hrs of Visit 1>*: X

### Completed by Subject

- Consent Form: X
- Demographic information: X
- NCCN/FACT‡ Lung Symptom Index-17 (FLSI-17): X
- NSCLC-SAQ: X
- Patient Global Impression of Severity (PGIS): X
- Patient Global Impression of Change (PGIC): X

---

† Note: Although subjects were scheduled for follow-up appointments one week after their Visit 1, retests completed up to 10 days following Visit 1 were accepted as valid assessments.

‡ National Comprehensive Cancer Network® (NCCN) – Functional Assessment of Cancer Therapy (FACT).
NSCLC-SAQ Item Examples (Tablet)

How would you rate the worst pain in your chest over the last 7 days?
- No Pain at All
- Mild Pain
- Moderate Pain
- Severe Pain
- Very Severe Pain

How often did you feel short of breath during usual activities over the last 7 days?
- Never
- Rarely
- Sometimes
- Often
- Always
Study Sample

Key eligibility criteria:
• Subject was at least 18 years of age with diagnosis of Stage IIIB or IV NSCLC
• Subject was naïve to treatment for NSCLC or had recovered from any prior treatment-related toxicity/adverse events to Common Terminology Criteria for Adverse Events (CTCAE) v4.03 grade 1 (mild) or better

Recruitment quotas:
• ≥ 30% with a diagnosis of Stage IV NSCLC
• ≥ 30% treatment naïve at enrollment
• ≥ 10% with an ECOG PS of 2 or higher
• Between 40% and 60% with a diagnosis of COPD
Analytical Approach

- Item descriptives
- Item reduction statistics
- Internal consistency reliability by item (Cronbach’s alphas if items removed)
- Rasch measurement theory to examine item ordering and person-to-item distribution
- Exploratory factor analysis to determine scale structure
- Cronbach’s alpha for total score
- Intraclass correlation coefficients to assess test-retest reliability
- Convergent (associations with the FLSI-17 Disease-Related Symptoms Subscale) and known-groups approaches to assessing construct validity
Results - Sample

Gender

- Female, 57%
- Male, 43%

Race n (%)

- White 132 (86.8)
- Black/African American 12 (7.9)
- Asian 3 (2.0)
- Other 5 (3.3)

Marital Status n (%)

- Married/living as married 92 (60.5)
- Divorced 24 (15.8)
- Widowed 21 (13.8)
- Never married 103 (32.7)

Age

- Mean: 64.3
- StDev: 9.8
- Range: 41-85

Gender

- Male, 43%
- Female, 57%

Race

- White 132 (86.8)
- Black/African American 12 (7.9)
- Asian 3 (2.0)
- Other 5 (3.3)

Marital Status

- Married/living as married 92 (60.5)
- Divorced 24 (15.8)
- Widowed 21 (13.8)
- Never married 103 (32.7)

Age

- Mean: 64.3
- StDev: 9.8
- Range: 41-85
## Results - Sample

<table>
<thead>
<tr>
<th>NSCLC Stage, n (%)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>IIIb</td>
<td>26 (17.1)</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>126 (82.9)</td>
<td></td>
</tr>
</tbody>
</table>

### Years since NSCLC diagnosis

- Mean (SD): 1.1 (1.5)
- Median [Range]: 0.5 [0.0-9.6]

### Treatment Status, n (%)

- Naïve: 50 (32.9)
- 1<sup>st</sup> Line: 49 (32.2)
- 2<sup>nd</sup> Line: 26 (17.1)
- 3<sup>rd</sup> Line: 27 (17.8)

### ECOG Performance Status, n (%)

- 0: 49 (32.2)
- 1: 78 (51.3)
- 2: 25 (16.5)

### Clinical diagnosis of COPD, n (%)

- No: 87 (57.2)
- Yes: 65 (42.8)

- At least 30% with a diagnosis of Stage IV NSCLC
- At least 30% treatment naïve
- At least 30% with a diagnosis of COPD
- Between 40% and 60% with a diagnosis of COPD
- At least 10% with an ECOG PS of 2 or higher
How often did you tire easily...?

N | 152
Mean (SD) | 2.14 (1.07)
Median | 2.00
Range | 0-4
N (%) Ceiling | 12 (7.9)
N (%) Floor | 14 (9.2)
N (%) Missing | 0 (0.0)
Correlation with:
FLSI-17 DRS | 0.622**
How would you rate the pain in your chest...?

Item-to-item correlation

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Other pain: 0.455</td>
<td></td>
</tr>
</tbody>
</table>

Item-to-total correlation

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.413 (Total items)</td>
<td></td>
</tr>
<tr>
<td>0.453 (Scale)</td>
<td></td>
</tr>
</tbody>
</table>

N = 152
Mean (SD) = 0.84 (1.06)
Median = 0.00
Range = 0-4
N (%) Ceiling = 77 (50.7%)
N (%) Floor = 4 (2.6%)
N (%) Missing = 0 (0.0)
Correlation with:
FLSI-17 DRS = 0.483**

Patient Global Impression of Severity (PGIS)
How would you rate your coughing...?

<table>
<thead>
<tr>
<th>N</th>
<th>152</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (SD)</td>
<td>1.05 (0.89)</td>
</tr>
<tr>
<td>Median</td>
<td>1.00</td>
</tr>
<tr>
<td>Range</td>
<td>0-4</td>
</tr>
<tr>
<td>N (%) Ceiling</td>
<td>42 (27.6%)</td>
</tr>
<tr>
<td>N (%) Floor</td>
<td>2 (1.3%)</td>
</tr>
<tr>
<td>N (%) Missing</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>Correlation with:</td>
<td></td>
</tr>
<tr>
<td>FLSI-17 DRS</td>
<td>0.426**</td>
</tr>
</tbody>
</table>

**Item-to-item correlation**

None

**Item-to-total correlation**

- 0.412 (Total items)
- 0.411 (Scale)

**Patient Global Impression of Severity (PGIS)**

- Not severe
- Mildly severe
- Very severe
Results – Items by Mean Score

- Mean scores (Day 1) for the 7 items of the NSCLC-SAQ are shown above
- Subjects used the full range of responses (i.e., 0, 1, 2, 3, and 4)
- Subjects provided answers for all NSCLC-SAQ items (there were no missing data)
Results – RMT Threshold Map
Results – Factor Analysis (All 7 items)

![Scree Plot]

<table>
<thead>
<tr>
<th>Component</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>6. tire easily</td>
<td>.853</td>
<td>.150</td>
</tr>
<tr>
<td>5. low energy</td>
<td>.849</td>
<td>.160</td>
</tr>
<tr>
<td>4. feel short of breath</td>
<td>.747</td>
<td>-.019</td>
</tr>
<tr>
<td>7. have a poor appetite</td>
<td>.618</td>
<td>.373</td>
</tr>
<tr>
<td>1. coughing at its worst</td>
<td>.507</td>
<td>.210</td>
</tr>
<tr>
<td>2. worst pain in your chest</td>
<td>.146</td>
<td>.835</td>
</tr>
<tr>
<td>3. worst pain in areas other than your chest</td>
<td>.136</td>
<td>.819</td>
</tr>
</tbody>
</table>

Note: Principal Component Analysis.
## Results – Item Correlation Matrix

<table>
<thead>
<tr>
<th></th>
<th>1. coughing</th>
<th>2. pain in chest</th>
<th>3. other pain</th>
<th>4. short of breath</th>
<th>5. low energy</th>
<th>6. tire easily</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. pain in chest</td>
<td>.297**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. other pain</td>
<td>.171*</td>
<td>.455**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. short of breath</td>
<td>.410**</td>
<td>.152</td>
<td>.136</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. low energy</td>
<td>.294**</td>
<td>.173*</td>
<td>.326**</td>
<td>.460**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. tire easily</td>
<td>.251**</td>
<td>.216**</td>
<td>.283**</td>
<td>.457**</td>
<td>.844**</td>
<td></td>
</tr>
<tr>
<td>7. poor appetite</td>
<td>.383**</td>
<td>.326**</td>
<td>.303**</td>
<td>.382**</td>
<td>.481**</td>
<td>.458**</td>
</tr>
</tbody>
</table>
The two pairs of items representing pain and fatigue were examined more closely. The decision was made to combine the two items for each concept in the provisional scoring algorithm as follows:

- **Fatigue**: The two items will remain in the questionnaire as the qualitative results warrant their inclusion; however, given the high correlation between the two items (0.84), indicating considerable conceptual redundancy, a score will be derived by taking the mean of the two items, thus becoming a single fatigue score.

- **Pain**: The two items will remain in the questionnaire as they are conceptually different (correlation of 0.46). Since the goal of the *NSCLC-SAQ* is to assess worst pain, wherever it manifests, a score will be derived by taking the most severe answer to either of the items, becoming a single pain score.
Results – Factor Analysis (All 7 items)

<table>
<thead>
<tr>
<th>Component 1</th>
<th>Item Description</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fatigue Score (Mean of Item 5 and Item 6)</td>
<td></td>
<td>.771</td>
</tr>
<tr>
<td>7. poor appetite</td>
<td></td>
<td>.768</td>
</tr>
<tr>
<td>4. feel short of breath</td>
<td></td>
<td>.712</td>
</tr>
<tr>
<td>1. coughing at its worst</td>
<td></td>
<td>.628</td>
</tr>
<tr>
<td>Pain score (Worst of Item 2 or Item 3)</td>
<td></td>
<td>.553</td>
</tr>
</tbody>
</table>

Note: Principal Component Analysis.
<table>
<thead>
<tr>
<th>Domain</th>
<th>Item</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cough</td>
<td>1. coughing at its worst</td>
<td>0, 1, 2, 3, 4</td>
</tr>
<tr>
<td>Pain</td>
<td>2. worst pain in your chest</td>
<td><em>Create a single score by selecting the highest severity (i.e., value) on either item</em></td>
</tr>
<tr>
<td></td>
<td>3. worst pain in areas other than your chest</td>
<td>0, 1, 2, 3, 4</td>
</tr>
<tr>
<td>Dyspnea</td>
<td>4. short of breath</td>
<td>0, 1, 2, 3, 4</td>
</tr>
<tr>
<td>Fatigue</td>
<td>5. low energy</td>
<td><em>Create a single score by calculating the mean of these 2 items</em></td>
</tr>
<tr>
<td></td>
<td>6. tire easily</td>
<td>0, 1, 2, 3, 4</td>
</tr>
<tr>
<td>Appetite</td>
<td>7. poor appetite</td>
<td>0, 1, 2, 3, 4</td>
</tr>
<tr>
<td><strong>NSCLC-SAQ Total Score (Sum the 5 domains)</strong></td>
<td><strong>Range 0 to 20</strong></td>
<td></td>
</tr>
</tbody>
</table>
NSCLC-SAQ covers five domains: Cough, Pain, Dyspnea, Fatigue, Appetite

All five of these domains must be non-missing to compute a total score.

**• Scoring algorithm**

- **Cough Domain Score**: score of the cough item, or missing if skipped
- **Fatigue Domain Score**: if both items present, compute mean; or use score from 1 item if the other is missing; or set to missing if both are skipped
- **Pain Domain Score**: if both items present, use most severe of both; or use score from 1 item if the other is missing; or set to missing if both are skipped
- **Dyspnea Domain Score**: score of the shortness of breath item, or missing if skipped
- **Appetite Domain Score**: score of the poor appetite item, or missing if skipped
- **NSCLC-SAQ Total Score**: sum all five domain scores; if any are missing, a total score is not computed
## Results – Convergent Validity (FLSI-17)

<table>
<thead>
<tr>
<th>NSCLC-SAQ Item/Score</th>
<th>Correlation (Pearson’s) With FLSI-17 DRS*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Coughing at its worst</td>
<td>0.421</td>
</tr>
<tr>
<td>2. Worst pain in your chest</td>
<td>0.486</td>
</tr>
<tr>
<td>3. Worst pain in areas other than chest</td>
<td>0.536</td>
</tr>
<tr>
<td>2/3. SAQ Pain Score</td>
<td>0.594</td>
</tr>
<tr>
<td>4. Short of breath</td>
<td>0.571</td>
</tr>
<tr>
<td>5. Low energy</td>
<td>0.680</td>
</tr>
<tr>
<td>6. Tire easily</td>
<td>0.692</td>
</tr>
<tr>
<td>5/6. SAQ Fatigue Score</td>
<td>0.713</td>
</tr>
<tr>
<td>7. Poor appetite</td>
<td>0.704</td>
</tr>
<tr>
<td>NSCLC-SAQ (5)</td>
<td>0.882</td>
</tr>
</tbody>
</table>

* FLSI-17 Disease-Related Symptoms Subscale-Physical

All correlations significant at the 0.01 level (2-tailed).
Results – Reproducibility

\[ \text{ICC} = 0.865 \ (0.802 - 0.909) \]

\[ \text{ICC} = 0.821 \ (0.756 - 0.930) \]

How would you rate your symptoms of your lung cancer at this time?

0  Not severe
1  Mildly severe
2  Moderately severe
3  Very severe
4  Extremely severe

PGIS: How would you rate your symptoms of your lung cancer at this time?

- Improved
- Worsened

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid</td>
<td>-4.00</td>
<td>.7</td>
<td>.7</td>
</tr>
<tr>
<td></td>
<td>-3.00</td>
<td>.7</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>-2.00</td>
<td>3.3</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td>-1.00</td>
<td>10.5</td>
<td>15.5</td>
</tr>
<tr>
<td></td>
<td>0.00</td>
<td>59.2</td>
<td>76.4</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>17.8</td>
<td>94.6</td>
</tr>
<tr>
<td></td>
<td>2.00</td>
<td>4.6</td>
<td>99.3</td>
</tr>
<tr>
<td></td>
<td>3.00</td>
<td>.7</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>148</td>
<td>97.4</td>
<td>100.0</td>
</tr>
<tr>
<td>Missing System</td>
<td>4</td>
<td>2.6</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>152</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>
Results – Known-Groups Validity

**Patient Global Impression of Severity**

- Not severe: n=72
- Mildly severe: n=41
- Moderately severe: n=31
- Very severe: n=5
- Extremely severe: n=3

**ECOG Performance Status**

- ECOG 0: n=49
- ECOG 1: n=78
- ECOG 2: n=25

**Self-reported Health Status**

- Excellent: n=7
- Very Good: n=41
- Good: n=51
- Fair: n=24
- Poor: n=29

Note: Overall significance p<0.001; post-hoc pairwise statistics show non-significance within comparisons between “Moderately severe,” “Very severe,” and “Extremely severe” (most likely due to small numbers).

Note: Overall significance p<0.001; post-hoc pairwise statistics show non-significance within comparisons between “Good” and “Fair” (p=0.35).
Results – Known-Groups Validity

Gender
p=0.622

Age
p=0.155
Conclusion

- The NSCLC-SAQ has been developed in accordance with FDA’s PRO Guidance.
- This study generated quantitative evidence of good item and scale performance and provided support for the content and construct validity of the data collected with the NSCLC-SAQ.
- The NSCLC-SAQ is currently included in longitudinal clinical trials to provide data for evaluation of additional measurement properties, including responsiveness.
For Further Information...

The initial quantitative results have recently been presented and published in Journal of Thoracic Oncology:

Challenges at the Site Level

Thomas Atkinson, Head, Behavioral Research Methods Core Facility, MSKCC
Recruitment challenges

• Delays in IRB approval at many sites
• Early stage and treatment-naïve patients had to be approached within hours (minutes) of being diagnosed with NSCLC
• Advanced stage patients (i.e., Stage IV)
  • Most not treatment naïve
  • Clinicians would not recommend patient participation
• Many patients not willing to come in during the 7-10 day window for the quantitative study when a clinic visit was not scheduled
Recruitment challenges

- Many specific COPD/stage/treatment naïve/ECOG PS quotas were extremely rare
  - Treatment naïve Stage IV patients with COPD
  - Stage IIIB patients
  - Treatment naïve patients with ECOG PS 2
  - Many ECOG PS 2 patients too sick to return for the follow-up quantitative assessment
  - Several instances of these patients being identified, but clinician declined to approach patients for participation
Recruitment challenges

• Solutions to meet challenges
  • Modified definition of “treatment naïve” to include patients “naïve to chemotherapy at the subject’s current stage of NSCLC, or no chemotherapy in the last 6 months as of enrollment”
  • Opening of second New York site to fill specific quotas (i.e., ECOG PS 2, treatment naïve)
Summary of Regulatory Interactions

Kendra DeBusk, Principal Outcomes Research Scientist, Genentech, Inc.
Regulatory Interactions

• Scoping Stage Summary Document (May 2010 – July 2011)
  • Written feedback
  • Topics included draft conceptual framework
  • Face-to-face Type C meeting with FDA

• Concept Elicitation Protocol and Study Documents (May-October 2013)

• Qualitative Research Summary (June-December 2014)
  • Written feedback
  • Topics included clarification of patient population and item/domain wording

• Teleconference with FDA (June 2015)
  • Topics included pain, fatigue, appetite items; recall/response; and use of single item for both pain and fatigue

• Quantitative Protocol and SAP (June-October 2015)
  • Written feedback
  • Topics included information on patients (treatment naïve and advanced stage)

• Qualification Briefing Package (March 2017)
  • Included Quantitative Pilot Study Report prepared in Fall 2016; not submitted separately
Ongoing and Future Activities

• Dissemination
  • Summary of qualitative research has been published
  • Summary of quantitative research has been presented at WCLC 2016
  • Once qualified for exploratory use
    • Manuscript describing quantitative testing
    • Presentation of FDA qualification process for the NSCLC-SAQ at scientific meetings

• Ongoing activities
  • Measure currently being used by WG sponsors in their clinical trials
  • Over 50 translations have been prepared for the NSCLC-SAQ
  • Inclusion in WG sponsor’s trials to allow for evaluation of longitudinal measurement properties
Panel Discussion and Q & A

Moderator
– Kendra DeBusk, PhD – Principal Outcomes Research Scientist, Genentech, Inc.

Presenters
– Astra Liepa, PharmD – Principal Research Scientist, Eli Lilly and Company
– Kelly McCarrier, PhD, MPH – Senior Research Scientist, Health Research Associates, Inc.
– Don Bushnell, MA – Associate Director, Health Research Associates, Inc.
– Thomas Atkinson, PhD – Head, Behavioral Research Methods Core Facility, MSKCC
– Kendra DeBusk, PhD – Principal Outcomes Research Scientist, Genentech, Inc.

Panelists
– Paul Kluetz, MD – Acting Deputy Director, Office of Hematology and Oncology Products, FDA
– Stephen Joel Coons, PhD – Executive Director, PRO Consortium, C-Path